Engine combustion modeling method based on hybrid drive

Deng Hu, Hechun Wang [*], Chuanlei Yang, Binbin Wang, Baoyin Duan, Yinyan Wang

College of Power and Energy Engineering, Harbin Engineering University, 150001, Harbin, China

ARTICLE INFO

ABSTRACT

Accurate and comprehensive reconstruction of in-cylinder combustion process is essential for timely monitoring of engine combustion state. This article developed a method based on the zero-dimensional (0-D) physical model integrated with big data. The traditional 0-D prediction model based on cumulative fuel mass is improved, the factor of in-cylinder temperature is introduced to adjust the heat release rate, which solves the problem of difficulty in calibrating the heat release rate. Then, convolutional neural network-gated recurrent unit (CNN-GRU), as a deep neural network, including a special convolutional layer and a gated recurrent unit (GRU) neural network is designed for the parameters to be calibrated in the model. The 0-D predictive combustion model is constructed by combining the physical model with CNN-GRU, the combustion process is simplified and reconstructed. The fitting results show that the 0-D physical model based on improved cumulative fuel mass approach is an effective method to reflect the heat release law. Under non-calibration conditions, the root mean square error (RMSE) value of peak firing pressure (PFP) based on CNN-GRU prediction model is 0.5862. The prediction model is a promising method to realize online fitting and optimization of combustion process.

1. Introduction

Recently, the demand for energy saving and emission reduction in marine engines, the real-time control of the combustion process in the engine cylinder becomes crucial [1,2]. The rapid increase in the calculation capacity of the electronic control unit (ECU) makes it possible for the engine to achieve complex control of in-cylinder combustion [3]. Therefore, the real-time control of the combustion process has become a research hotspot at home and abroad. This study seeks to reduce emissions, and optimize the performance of the engine. By establishing a combustion model, the combustion process can be better understood and controlled, thereby improving the combustion efficiency of the engine.

Currently, real-time control of the combustion process is mainly implemented in two modes: model-driven control and data-driven control. Model-driven controls are primarily which employs cylinde r pressure signal and map interpolation. ss [4] is based on the engine equipped with a large number of expensive equipment such as cylinder pressure sensors and signal amplifiers. Because these devices are currently difficult to be equipped on mass production engines, this technology is mainly used in the test stage and cannot be promoted on a large scale. The control based on map interpolation [5] can achieve accurate control of the combustion process when there are few control variables. However, as the variables increase, the test data required for map calibration also increases, and the control speed and accuracy of the model begins to decline sharply. A data-driven method is one that uses heuristic rules to directly
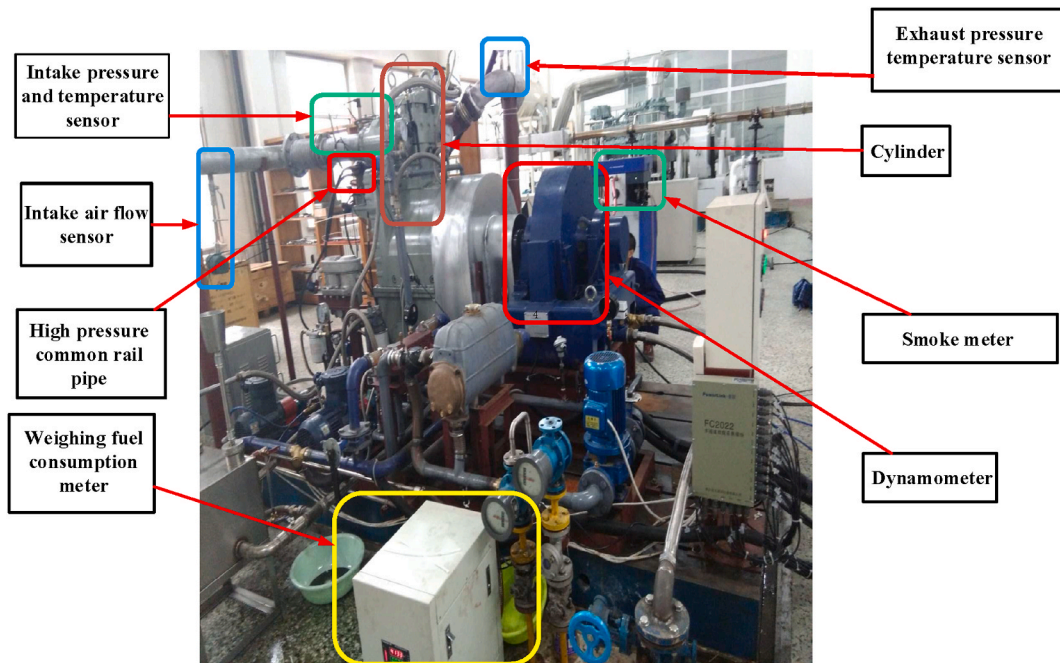
Fig. 1. The main layout of TBD620 diesel engine test bench.

establish the relationship between operating parameters and combustion results from the monitored test data. The data-driven combustion process control doesn't take physical interpretability into account and therefore does not need to construct a complete prediction model. Instead, it conducts big data analysis on the collected historical information, and uses artificial intelligence methods such as machine learning to train and reconstruct the model [6–9]. Cheng Ma used extensive experimental data of a dual-fuel engine to construct a 0-D prediction model through long short-term memory (LSTM), which achieved a quickly response of operating parameters to fuel consumption and emissions. Wen Bin constructed the relationship between operating parameters and combustion results (indicated mean effective pressure (IMEP) and crank angle at which 50 % of the fuel mass fraction has burned (MFB50)) through the back propagation neural network. Research on data-driven reconstruction of combustion process is proceeding with full vigor and data-driven is more suitable for today's complex control systems with diverse functions and uncertainties. However, the huge amount of data required for model construction and the poor reliability of data have been important constraints on the development of combustion process control.

The hybrid-driven method integrates mechanistic and data-driven approaches [10–12], where the main idea is to summarize the uncertain parameters that vary with the environment in the process of constructing a physical model, and use the monitored big data to fit these parameters, so that the model can adapt to the accuracy requirements of the reconstructed combustion process under different working conditions. For example, in the process of constructing a combustion model based on Wiebe control, Hu Song [13] used algebraic analysis-least square method to calculate the parameters. The model combines the advantages of the algorithm and physical model to improve the computational speed while effectively explaining the variation of key parameters such as fuel injection law, in-cylinder temperature, heat release rate and combustion starting point during the combustion process, which greatly improves the interpretability of the model. Currently, there are relatively few engine modeling methods driven by the integration of mechanism and data, and related research is still in its infancy and promising.

At present, the physical model that can meet the real-time control is mainly 0-D model. There are two main modeling approaches to 0-D combustion model commonly used in the published literature for the heat release law of in-cylinder fuel. One is the combustion rule method based on mathematical equations, including EF method [14,15] and well-known Wiebe function [16,17]. The other is the cumulative fuel mass approach based on the concept of fuel injection [18]. The Wiebe function exerts excellent simulation capability for in-cylinder combustion process under rated or high load conditions, but in the case of multiple injections and multiple combustion phases coexisted, which need multiple Wiebe functions to fitting. Franz G. Chmela [19] assumed that the heat release rate was proportional to the mass of available fuel in the cylinder, and proposed a combustion model based on cumulative fuel mass, which was later improved by several scholars. Finesso added a function term to the model based on the results of Franz G. Chmela. This function term represents the influence of turbulent kinetic energy of fuel injection atomization, thus improving the simulation accuracy of heat release rate in the initial stage of combustion. And the combustion model based on cumulative fuel mass requires fewer calibration model parameters and has high fitting accuracy, which is suitable for multiple injection engines of various models. However, there are still large errors in the fitting process of the heating rate in the single injection, especially in the diffusion combustion stage, and new adjustment methods need to be introduced at this stage.

Based on the parameter fitting of data-driven combustion models, many researchers have used classical algorithms [20–24] for
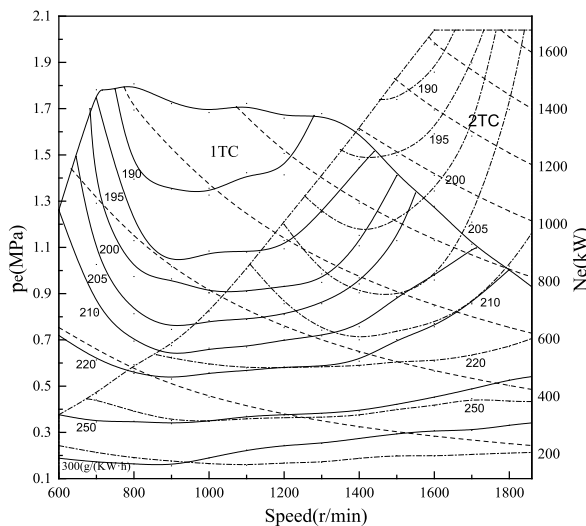
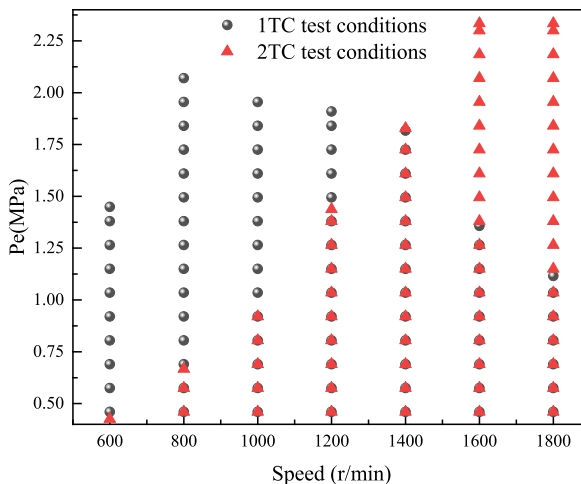**Fig. 2.** Universal characteristic diagram of the TBD620V12.

Fig. 3. Test conditions of TBD620 single-cylinder engine.

calculation. In recent years, with the rapid development of deep learning algorithms, deep learning algorithms are rapidly applied to the field of engine modeling [25,26]. Shahid [27] used convolutional neural network (CNN) to establish the variation of engine operating parameters with misfire and load by analyzing the in-cylinder combustion process, which in turn created a real-time diagnostic system for detecting misfire and load conditions. Alcan [28] developed a real-time prediction model for diesel soot emissions using the GRU neural network. Shen et al. [29] constructed a NOx emission model by combining CNN with LSTM, which greatly improved the prediction accuracy. GRU is simpler and easier to train than LSTM structure while solving the gradient problem during long term memory and back propagation.

Song Hu [30] used the traditional 0-D physical model based on cumulative fuel mass combined with artificial intelligence neural network (ANN) and empirical formula (EF) to obtain a control-oriented rapid response model. Based on the improved 0-D physical model of cumulative fuel mass, the 0-D prediction model is constructed by combining CNN-GRU deep learning neural network. The construction of combustion model based on hybrid drive can realize the real-time prediction of combustion process.

2. Diesel engine test bench and test operating point

The current work takes the single-cylinder diesel engine (HeChai Group Co., model TBD620) as the research object, developed by MWM (Motorenwerke Mannheim), Federal Germany, as the test platform. The main layout of TBD620 diesel engine test bench as shown in Fig. 1. Because the single-cylinder engine does not match the supercharger, the load characteristics of the engine cannot be determined directly. Therefore, the test operating point is based on the universal characteristics of TBD620V12, and the single-cylinder

Table 1

Comparison of main parameters between TBD620V12 and the single-cylinder engine.

Serial No.	Parameters	TBD620V12	TBD620 single-cylinder engine
1	Number of cylinders	12	1
2	Rated power (kW)	1680	146
3	Rated speed (r/min)	1800	1800
4	Fuel system	high pressure common rail	Mechanical fuel injection system
5	Cylinder diameter (mm)	170	170
6	Piston path (mm)	195	195
7	Compression ratio (mm)	13.5	13.5
8	Fuel consumption rate (g/(kW·h))	198.2	225
9	Combustion mode	direct injection	direct injection

Table 2

Main parameters of test equipment.

Equipment names	Model
Dynamometer	GW320
Combustion analyzer	AVL INDIMDUL 621
Pressure sensor	KISTLER 6058A
Fuel consumption meter	FC2210

engine is equipped with a blower at the air inlet to test the working conditions. Fig. 2 shows the universal characteristic diagram, from which we can observed that the TBD620V12 is supercharged by sequential turbocharging (STC). Therefore, the TBD620 single-cylinder engine tests the working conditions by changing the blowing volume of the blower, and the specific working points are shown in Fig. 3. Test points for steady-state testing, consisting of two parts: 83 test points selected under the 1 TC test conditions and 68 test points selected under the 2 TC test conditions, for a total of 151 test points. Due to the difficulty of testing the test points, the linear interpolation of the working points is carried out on the basis of the test points. The test points plus interpolation points have a total of 590 working points.

Table 1 shows the comparison of parameters between TBD620V12 and single cylinder engine. Both of them are four-stroke, water-cooled marine diesel engines, corresponding to the same rated speed, cylinder diameter, stroke and compression ratio, but there are some differences in fuel system, single-cylinder calibration power, effective fuel consumption rate and fuel system. Table 2 shows the main test equipment. GW320 Eddy Current Dynamometer for Measuring Engine Power, a KISTLER 6058A pressure sensor is used to measure in-cylinder pressure, an AVL INDIMDUL 621 combustion analyzer is used to analyze the combustion process.

3. Construction of 0-D prediction model based on cumulative fuel mass approach

3.1. Effect of in-cylinder temperature on heat release rate

For simulation of combustion processes, the number of Wiebe functions to be selected for a single combustion cycle usually increases exponentially due to the increase in the number of injections, making it more difficult to calibrate. Therefore, the cumulative fuel mass approach is used to solve the problem, as Eq. (1):

$$\begin{cases} \dfrac{dQ_{ch,pil,j}}{dt}(t) = k_{pil,j} \times \left[Q_{fuel,pil,j}(t-\tau) - Q_{ch,pil,j}(t) \right] \\[2mm] \dfrac{dQ_{ch,main}}{dt}(t) = k_{1,main} \times \left[Q_{fuel,main}(t-\tau_{main}) - Q_{ch,main}(t) \right] + k_{2,main} \times \dfrac{dQ_{fuel,main}(t-\tau_{main})}{dt} \\[2mm] Q_{fuel,j} = \displaystyle\int_{t_{SOI}}^{t} \dot{m}_{f,inj}(t)HLdt t \le \tau_{EOI} Q_{fuel,j} = \int_{t_{SOI}}^{t} \dot{m}_{f,inj}(t)HLdt t > \tau_{EOI} Q_{ch} = \sum_{j=1}^{n} Q_{ch,j} \end{cases} \tag{1}$$

Where HL is low heating value of fuel, pil is pilot injection, $main$ is main injection, Q_{ch} is chemical heat release, τ is ignition delay coefficient, k_{pil} is combustion parameters corresponding to pilot injection combustion, $k_{1,main}$ and $k_{2,main}$ are the combustion parameters corresponding to the main injection combustion, $\dot{m}_{f,inj}$ is fuel injection rate and Q_{fuel} is chemical energy.
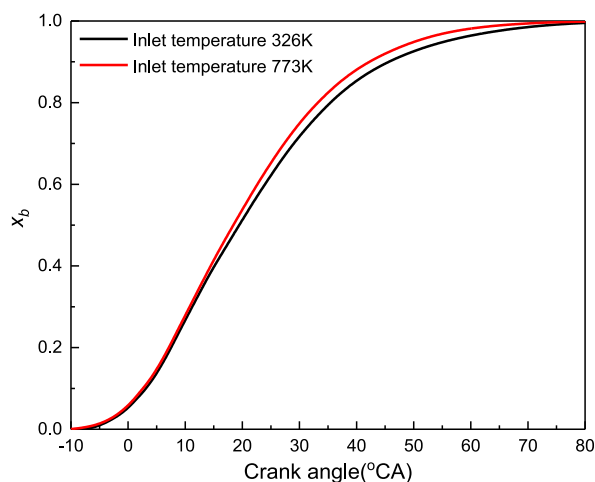
Fig. 4. Combustion chamber model.

Fig. 5. Effect of in-cylinder temperature on heat release rate curve.

For the combustion process of a multi-k injection engine, both Wiebe function and cumulative fuel mass approach can be used for fitting. However, in the case of the coexistence of premixed and diffusion combustion phases, the number of parameters to be calibrated for the combustion process increases exponentially if double Wiebe function is adopted. When the fuel system is in single injection mode, the combustion model based on the cumulative fuel mass is also applicable, and only three calibration parameters are required. There are various ways of influencing fuel burn rate, of which the factors affecting heat release rate curve of the combustion process considered in Eq. (1) are mainly the turbulent kinetic energy and the fuel injection law generated by fuel injection atomization. However, in the actual combustion process, the fuel has been undergoing complex physical and chemical changes since it was injected into the cylinder, and these changes are closely linked to the transfer of heat, which is achieved in three ways. The heat convection in the combustion process is mainly affected by turbulence, which has been reflected in Eq. (1). The heat conduction is mainly related to the temperature difference between the mediums. The initial temperature is about 333 K, while the maximum temperature in the cylinder is more than 2000 K, so there is a huge difference during the fuel and cylinder temperature. The heat radiation in the engine cylinder is mainly caused by high temperature gas and carbon particles, which is affected by the temperature in the cylinder. Ref. [30] made a sensitivity analysis of the possible factors affecting the heat release rate. The fuel injection quantity, engine speed, fuel injection timing, fuel injection temperature, and air temperature of the intake pipe all have a great influence on the heat release rate. Therefore, which is necessary to quantify the cylinder temperature on the heat release rate.

Fig. 4 shows the combustion chamber geometry used in the simulation process. CONVERGE is a new generation of thermal fluid analysis software developed by Convergent Science Company in the United States. Through accurate, efficient and reliable three-

Table 3
In-cylinder temperature corresponding to different crank angles.

Crank angle(°CA)	326K	773K
720°CA		
755°CA		
Compression starting point		
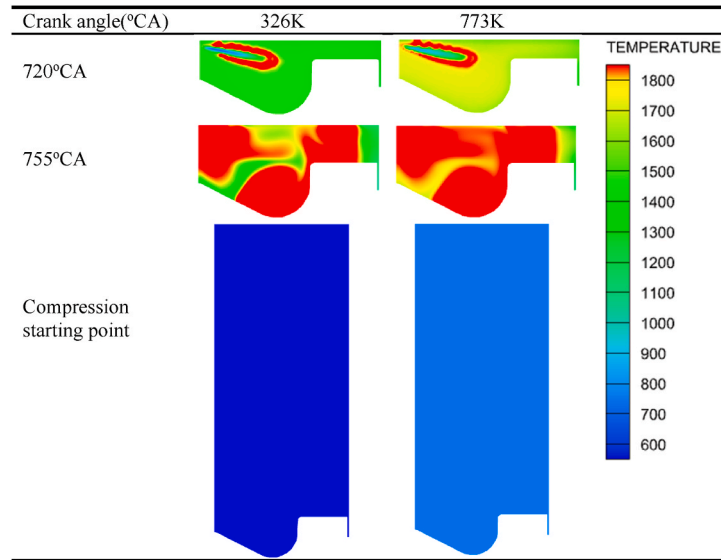
dimensional fluid analysis process, detailed simulation of engine combustion process can be realized. To study the influence of in-cylinder temperature variation on heat release rate curve, based on the 3.0 version of Converge software and the three-dimensional prediction model that has been calibrated [31], under rated conditions, the intake flowrate and injection volumes remain unchanged, and the intake air temperature is set to 326 K (the original machine intake air temperature) and 773 K, respectively, in order to roughly simulate the difference in in-cylinder temperature, and compare the final results. The reduction of intake flow rate caused by the sharp increase in intake temperature is offset by increasing the intake pressure. The intake pressure is set to 0.366 MPa (the original intake pressure) and 0.666 MPa, respectively.

Fig. 5 shows the effect of different in-cylinder temperatures on the heat release rate curve. The difference of temperature is realized by adjusting the intake temperature, which in turn affects the actual heat release rate curve. The moment of fuel ignition with a higher in-cylinder temperature is 3°CA earlier than that of the original engine. This is because the increase of in-cylinder temperature greatly reduces the time required for fuel vaporization under the condition of constant fuel injection temperature. However, the increase of in-cylinder temperature is more reflected in the diffusion combustion stage, especially at 35°CA after the top dead center (TDC), when cumulative heat release rate error caused by the temperature is more than 3 %. The high temperature in the cylinder not only transfers heat to the cylinder liner, piston and cylinder head, but also transfers heat to the oil-air mixture that is reacting or not reacting. Table 3 shows the in-cylinder temperature slices corresponding to different crank angles at the initial in-cylinder temperatures of 326 K and 773 K, respectively. As can be observed from Table 3, the in-cylinder temperature is the same as the intake air temperature at the moment of the start of compression, proving that the differential design of in-cylinder temperature can be realized by changing the intake temperature without changing the intake flow rate. There is a significant difference in in-cylinder temperature between TDC and 35°CA after TDC. Among them, at TDC, the increase of in-cylinder temperature has a great effect on the vaporization and decomposition of fuel spray. At 35°CA after TDC, the injection process is complete, and the influence of turbulent kinetic energy brought by injection atomization on oil-air mixing process is also reduced. At this time, the heat transfer of in-cylinder temperature to oil-air mixture can meet the activation energy required for chemical reaction of fuel. Therefore, study the influence of in-cylinder temperature on heat release rate is necessary when constructing a combustion model based on cumulative fuel mass approach.

3.2. Optimization of the model

Through the analysis of the three-dimensional combustion process, we can conclude that the heat release rate of in-cylinder combustion process is closely linked to the in-cylinder temperature. Based on the above analysis, for single injection engines, the author further improved Eq. (1) by adding a function term that is proportional to the increase in internal energy of the unburned accumulated fuel during fuel injection. Assuming that the fuel in the injection area is not mixed with air at the moment of combustion injection and that the fuel spray is rapidly heated to the in-cylinder temperature, the improved combustion model can be expressed as Eq. (2):
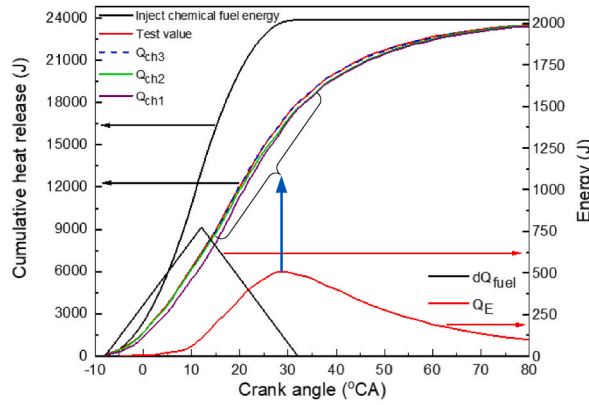
Fig. 6. Comparison of heat release rate curves of different combustion models.

$$\begin{cases} \dfrac{dQ_{ch}}{dt}(t) = k_1 \times \left[Q_{fuel}(t-\tau) - Q_{ch}(t) \right] + k_2 \times \dfrac{dQ_{fuel}(t-\tau)}{dt} + k_3 \times Q_E(t-\tau) \\[4mm] Q_{fuel} = \displaystyle\int_{t_{SOI}}^{t} \dot{m}_{f,inj}(t) HL dt t \le \tau_{EOI} \quad Q_{fuel} = \displaystyle\int_{t_{SOI}}^{t} \dot{m}_{f,inj}(t) HL dt t > \tau_{EOI} \quad Q_E = \left(\displaystyle\int_{t_{SOI}}^{t} \dot{m}_{f,inj}(t) - \dfrac{Q_{ch}(t)}{HL} \right)(T - T_0) \end{cases} \tag{2}$$

where Q_E is the increase in internal energy of the unburned accumulated fuel during fuel injection, T is in-cylinder temperature, and T_0 is fuel temperature at injection time.

Fig. 6 shows the comparison results of heat release rate curves of different combustion models based on cumulative fuel mass approach, where Q_{ch1} represents the burnt fraction curve without considering the influence of turbulent kinetic energy and in-cylinder temperature on the heat release rate, Q_{ch2} represents burnt fraction curve after only considering the influence of turbulent kinetic energy on heat release rate, Q_{ch3} represents the burnt fraction curve after considering the influence of in-cylinder temperature and turbulent kinetic energy. The influence of turbulent kinetic energy on heat release rate is mainly reflected in the initial combustion stage, where the turbulent kinetic energy generated by fuel injection atomization effectively improves the mixing effect of oil and gas and greatly promotes the combustion rate of fuel. The in-cylinder temperature can influence heat release rate mainly occurs in diffusion combustion stage, where the in-cylinder temperature reaches the peak value of the combustion process, and the heat absorbed per unit fuel before combustion also reaches its highest value. When the temperature is lower than 2500 °C, the effect of temperature on flame propagation speed increases geometrically, so that the in-cylinder flame propagation speed increases significantly as the injected fuel itself absorbs heat more rapidly. Although the Q_E derived from the above simplification cannot be taken as a change in the total energy of the fuel and is somewhat contrary to the physical meaning, it can effectively regulate the heat release rate, which is necessary for the accurate prediction of the heat release rate curve. In summary, the turbulent kinetic energy can effectively correct the heat release rate curve in the initial stage of combustion, and the increase of internal energy of unburned accumulated fuel can effectively adjust the heat release rate curve.

Fig. 7 is the comparison results of the heat release rate curves based on the cumulative fuel mass approach and the standard Wiebe function. The data in Fig. 8 were calculated by 3.9 version of Python, and 2022 version of Origin software was used to generate the graphs. It can be observed from the figure that under the four classical working conditions, the improved cumulative fuel mass approach has a higher fitting accuracy than the traditional cumulative fuel approach and the standard Wiebe function. As the load decreases, the difference value between the burnt fraction curves calculated by the three methods and the experimental value increases to varying degrees, indicating that the decrease in load will reduce the fitting accuracy of the three methods [32,33]. The main reason is that under high load conditions (Fig. 7(a), (b) and (c)), the combustion process is dominated by diffusion combustion and the combustion form is relatively simple, so it is less difficult to calculate. The continuous decrease of load caused a gradual increase of ignition delay period. The peak of the heat release rate curve gradually changes from a single peak to double peaks (Fig. 7(d)). The standard Wiebe function can only accurately calculate a single peak, so the calculation accuracy is significantly decreased under low load conditions. The cumulative fuel mass approach also has varying degrees of decrease in calculation accuracy under low load conditions, where the influence of the turbulent kinetic energy caused by in-cylinder temperature and injection is significantly weakened due to the reduced fuel injection quantity and the lower in-cylinder temperature, so the error curves corresponding to Q_{ch2}
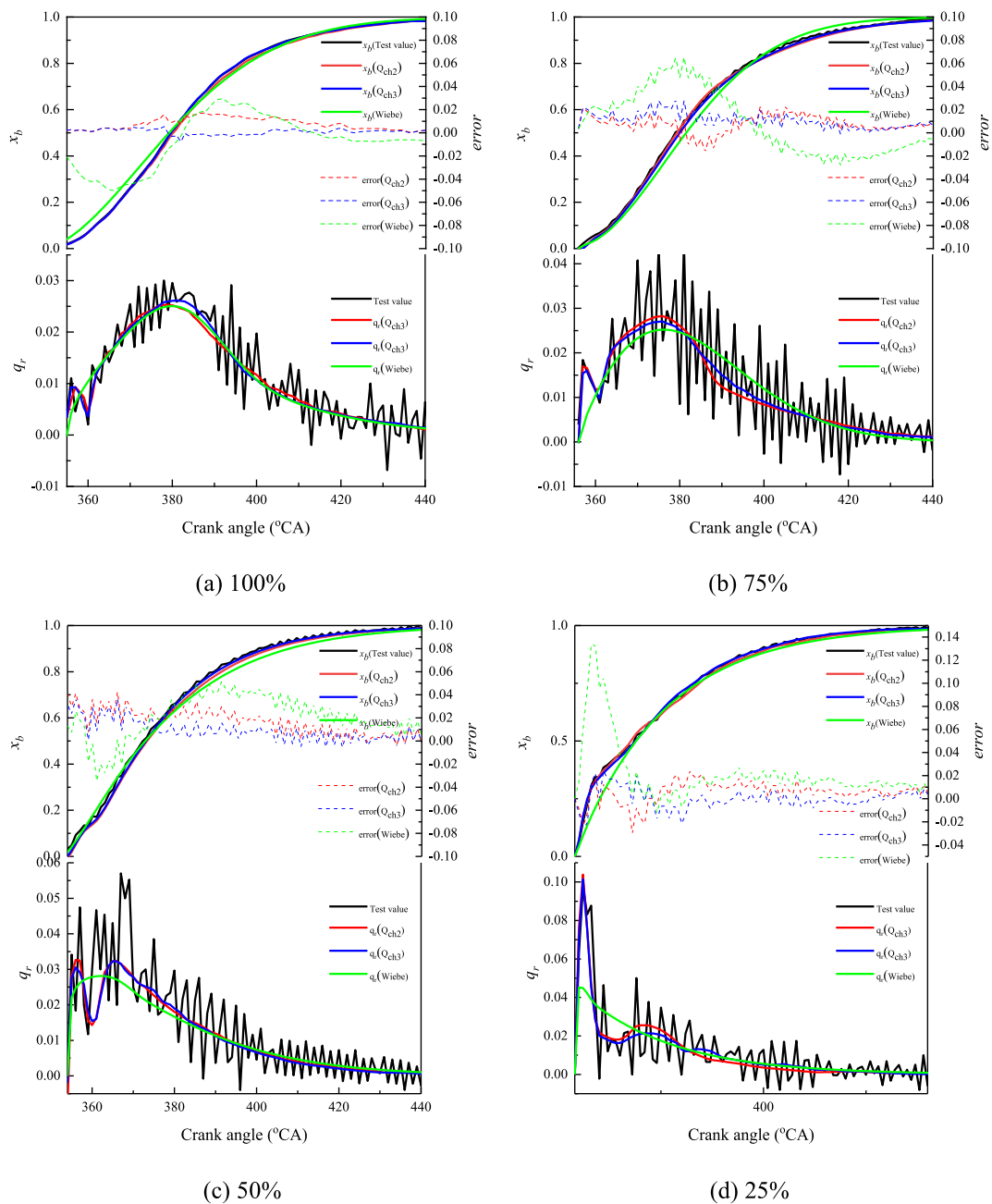
Fig. 7. The comparison results of the heat release rate curves calculated by different methods under 100 %, 75 %, 50 % and 25 % working conditions.

and Q_{ch3} are close. Compared with the traditional method, the improved cumulative fuel mass approach has a better regulation of the instantaneous heat release rate in the diffusion combustion stage. The traditional cumulative fuel mass approach needs to adjust k_1 and k_2 to achieve the final total heat release. Therefore, the instantaneous heat release rate of the traditional cumulative fuel mass approach in the premixed combustion stage and the tail combustion stage is lower than that of the improved cumulative fuel mass approach, and the corresponding calculation accuracy also decreases.
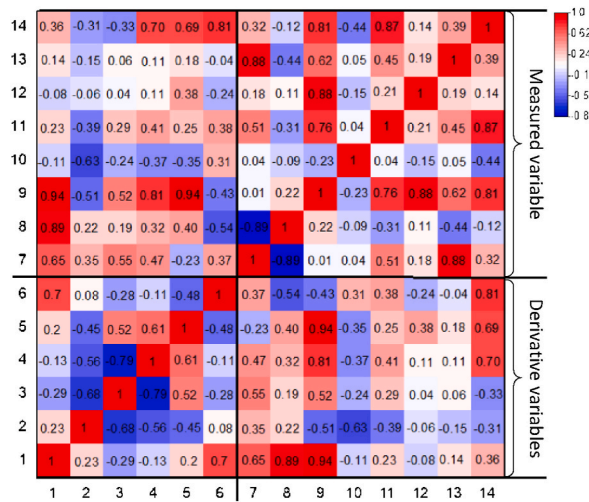
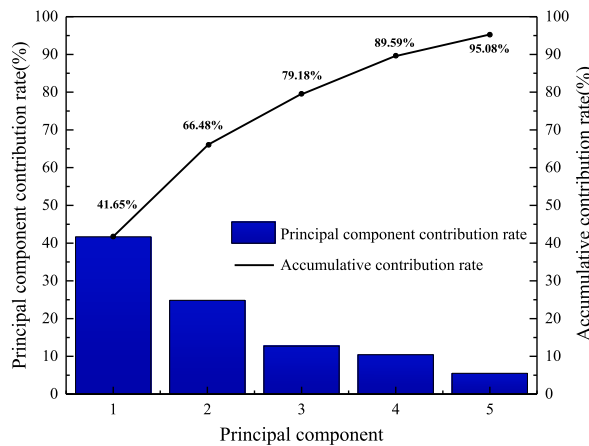**Fig. 8.** Pearson correlation coefficient between variables.

Fig. 9. Principal component analysis results.

4. Identification of mapping relationship between parameters based on data-driven

4.1. Determination of the mapping relationship between parameters in the model

The EF obtained by power function fitting shows the number of independent variables has no significant impact on fitting accuracy, and greatly increases the complexity of the formula. Similarly, the excessive increase in the number of independent variables may lead to overfitting of neural networks. Therefore, in order to determine a reasonable number of independent variables, sensitivity analysis of the parameters is necessary. Pearson correlation coefficient is one of the main methods to measure the correlation between variables [34], which can be expressed as Eq. (3):

$$P(x_i, x_j) = \frac{cov(x_i, x_j)}{\sqrt{var(x_i) \times var(x_j)}} \tag{3}$$

where $var(x_i)$ and $var(x_j)$ are variances of x_i and x_j respectively, $cov(x_i, x_j)$ is covariance of x_i and x_j.

Taking 151 sets of data under steady-state test conditions as the training set, calculate the Pearson correlation coefficient of the parameters, then select the candidate operation parameters corresponding to the parameters to be fitted. From Fig. 8, we can see the Pearson correlation coefficient heat map between the variables. The numbers 1–6 in the figure are τ, k_1, k_2, k_3, Q_{glob} and Q_{evape}, respectively, noted as derived variables, and numbers 7–14 are for engine speed n, start of injection SOI, fuel injection mass m_{mfuel}, intake manifold pressure P_{int}, exhaust manifold temperature T_{ex}, intake manifold temperature T_{int}, exhaust manifold pressure P_{ex}, and air mass flow m_{air}, respectively, noted as measured variables. The Pearson correlation coefficient takes values between −0.89 and 1.

Table 4
Variance contribution rate and CPV of each principal component.

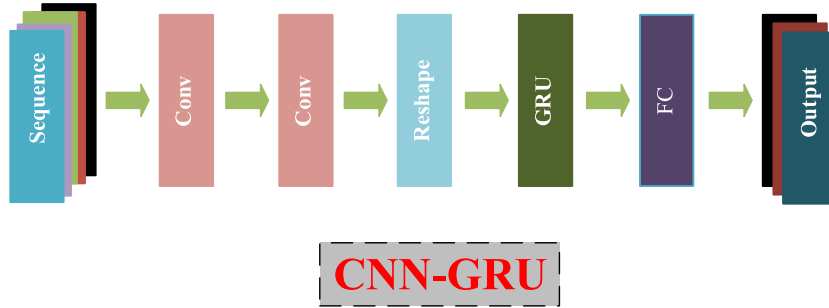| Principal components	Eigenvalue	Variance contribution rate (%)	CPV (%)
1	3.64	41.65	41.65
2	2.17	24.83	66.48
3	1.11	12.70	79.18
4	0.91	10.41	89.59
5	0.48	5.49	95.08
6	0.36	4.12	99.2
7	0.05	0.57	99.77
8	0.02	0.23	100

Fig. 10. CNN-GRU neural network structure diagram.

The size of the absolute value indicates the strength of the linear correlation, and a smaller absolute value indicates a weaker correlation between two parameters. From Fig. 8, the correlation coefficients between the six derived variables and the eight measured variables can be obtained, with T_{ex} and P_{ex} having the lowest influence on the derived variables. In the analysis of the two measured variables, m_{air} and T_{int}, it is found that T_{int} was chosen to replace m_{air} due to the lower stability and accuracy of the m_{air}, as well as the larger correlation coefficient with T_{int}, which had a strong physical correlation. For mass-produced engines, P_{ex} is usually not measured and has a strong correlation with engine speed, so P_{ex} is not considered as an input variable.

Pearson correlation coefficient only considers the correlation between two variables, and the interaction between multiple factors is not considered at the same time. On this basis, this paper realizes the extraction and dimension reduction of the feature of sample data by introducing principal component analysis (PC A) [35]. In Fig. 9, the covariance, eigenvalue and eigenvector required for the principal components are calculated in the same way as the Pearson correlation coefficient. The principal components of the above eight measured variables were selected by cumulative percent variance (CPV) [36]. The cumulative variance contribution rate of the first five components reached 95.08 %, which was greater than 95 %, meeting the accuracy requirements for nonlinear fitting. As can be observed from Table 4, the principal components with the lowest variance contribution rate are T_{ex}, P_{ex} and m_{air}. The independent variable parameters of the EF and the neural network are finally determined as n, SOI, m_{mfuel}, T_{int} and P_{int}, respectively.

4.2. Parameter relationship identification model based on EF

Through the sensitivity analysis performed, the independent variable set corresponding to the parameters to be calculated in the model is obtained, and then EF of each parameter is obtained by using the power function as primary function. Finally, the mapping relationship between the parameters to be calculated and the independent variable set in the combustion model is determined, as shown in Eqs. (4)–(9):

$$\tau_p = 6.69 \times 10^4 \cdot n^{0.593} SOI^{-1.25} m_{fuel}^{-0.38} P_{int}^{-0.41} T_{int}^{-0.23} \tag{4}$$

$$k_1 = 2.13 \times 10^6 \cdot n^{-1.1} SOI^{1.1} m_{fuel}^{-0.42} P_{int}^{-0.59} T_{int}^{-3.13} \tag{5}$$

$$k_2 = 1.06 \cdot n^{-0.15} SOI^{-0.26} m_{fuel}^{0.43} P_{int}^{0.38} T_{int}^{0.042} \tag{6}$$

$$k_3 = 11.3 \cdot n^{-0.436} SOI^{1.78} m_{fuel}^{-2.1} P_{int}^{-0.24} T_{int}^{0.13} \tag{7}$$

$$Q_{f,evap} = 6.18 \times 10^{-5} \cdot n^{0.342} SOI^{2.38} m_{fuel}^{-0.29} P_{int}^{-0.52} T_{int}^{0.023} \tag{8}$$

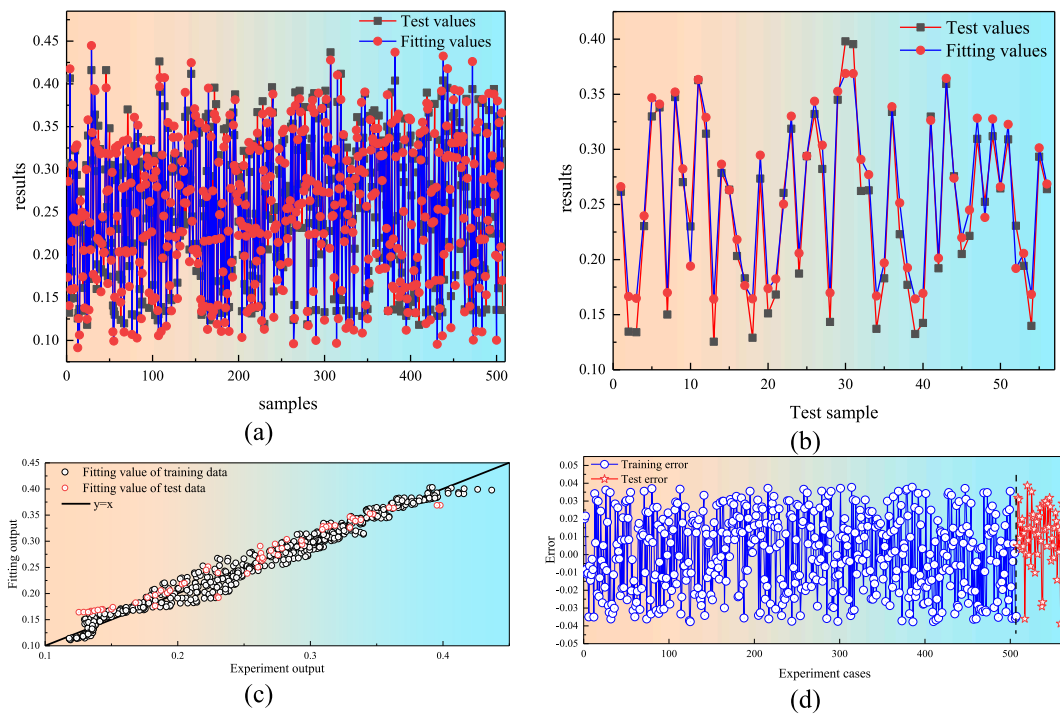$$Q_{ht,glob} = 1.74 \cdot n^{-0.53} SOI^{3.22} m_{fuel}^{-0.47} P_{int}^{0.32} T_{int}^{-1.87} \tag{9}$$

Fig. 11. The regression analysis and error distribution between the simulation results and the test results of the training set and the test set based on EF identification model.

Table 5
Statistical values of different neural network performance.

	EF		CNN-GRU	
	Training	Test	Training	Test
R^2	0.9513	0.9507	0.9861	0.9851
MAPE%	8.61 %	8.41 %	3.11 %	3.50 %
RMSE	0.0213	0.0196	0.0096	0.0110

4.3. Parameter relationship identification model based on deep learning

According to the fuel delivery parameters, the burnt fraction curve and the in-cylinder temperature obtained from the test, the parameters to be calculated in the improved cumulative fuel mass formula can be calculated. Taking the engine operating parameters as the input parameters and the parameters to be calculated as the output parameters, the identification model is constructed by using the neural network, and the nonlinear relationship between the operating parameters and the parameters to be calculated in the model is established.

To further improve the nonlinear mapping effect of engine operating parameters and combustion model parameters, the nonlinear fitting of the two is carried out by CNN-GRU deep learning neural network in this work. CNN can perform efficient extraction in the face of complex network structures. It was first designed for image data, but CNN is a deep neural network with a parameter sharing mechanism based on convolution, which can effectively extract features from data. Therefore, it can also adapt to the processing of one-dimensional data. The relationship between operating parameters and combustion model parameters is a multi-input and multi-output relationship. It is difficult to accurately obtain an explicit mathematical expression between variables. CNN can extract the overall characteristics of the data within a certain range from the input data as the input of the GRU model, it can greatly improve the recognition effect of GRU on the data. In Fig. 10, the input data is convolved twice to perform scanning local feature extraction on the data, passes through the flatten layer, and finally passes through the output layer.

4.4. The identification effect of different calculation methods on the mapping relationship between parameters

To study the identification effect of different solution methods on the mapping relationship between parameters, it is necessary to keep the sample order unchanged. 90 % of the samples were randomly selected as the training set, and the remaining 10 % of the samples were selected as the test set. In Fig. 11, the fitting results of the identification model based on EF to the parameters to be
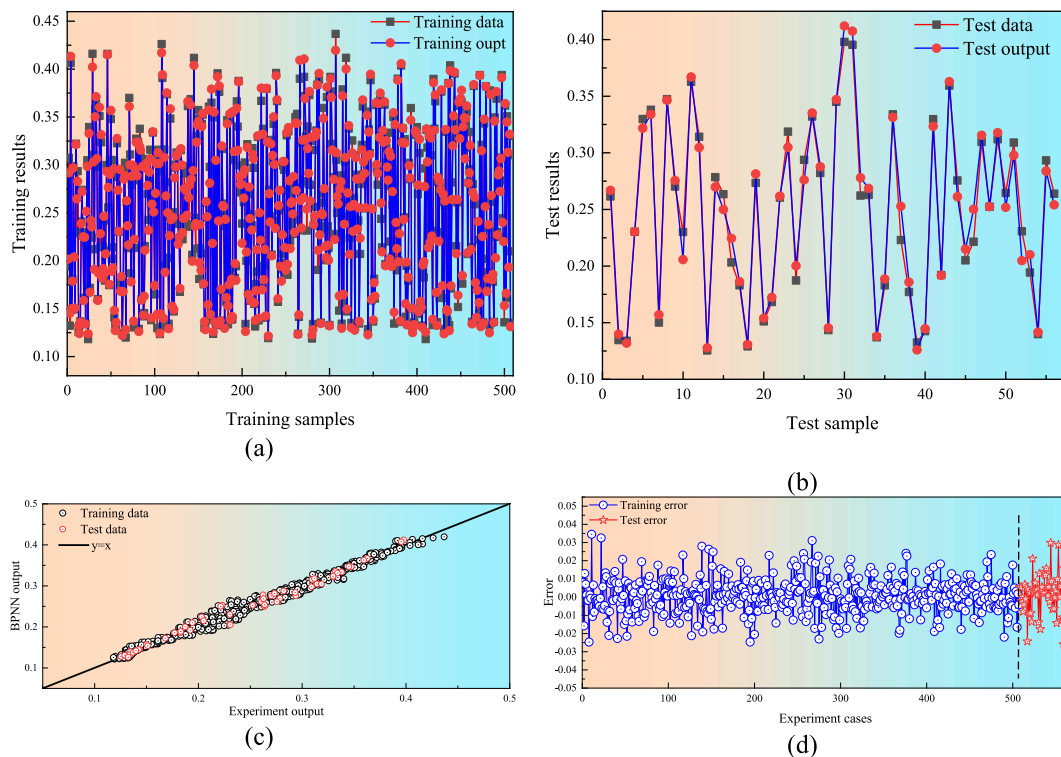
Fig. 12. The regression analysis and error distribution between the simulation results and the test results of the training set and the test set based on CNN-GRU identification model.

calculated. The figure shows the prediction results (Fig. 11(a) and (Fig. 11(b))) of the in-cylinder combustion prediction model based on the EF for k_2, and gives the regression analysis (Fig. 11(c)) and error distribution between the test results and the simulation results of the training set and the test set. From the figure we can see that, there is a large error between the tested and predicted values of k_2 (Fig. 11(d)), with the highest error exceeding 0.04. Meanwhile, from Table 5 (obtained using Python software), the coefficient of determination (R^2) of the training set and test set of k_2 are 0.9513 and 0.9507, respectively, which are lower than 0.97, and the regression effect is poor. Besides, the corresponding mean absolute percentage error (MAPE) values are 8.61 % and 8.41 %, respectively. In summary, although the prediction results based on EF can roughly reflect the variation law of the parameter values, the prediction accuracy cannot meet the minimum accuracy required by the model.

Fig. 13 is the prediction result (Fig. 12(a) and (Fig. 12(b))) of the identification model based on CNN-GRU for the parameters to be calculated, from which the prediction accuracies of k_2 (Fig. 12(c) and (Fig. 12(d))) have been greatly improved compared with the identification model based on EF. As can be observed from Table 5, the R^2 of the training set and the test set of the identification model based on CNN-GRU are 0.9861 and 0.9851, respectively, which are both greater than 0.97, with an increase of 3.48 % and 3.44 % respectively compared to EF. The MAPE value of the test set of the identification model based on CNN-GRU is 3.5 %, which is less than 5 % and is 58.38 % lower than the MAPE value of EF. In addition, the RMSE based on CNN-GRU of the test set of the identification model is 0.011, which is 0.086 less than EF. Therefore, we can recognize that identification effect of parameters by GRU neural network based on optimized convolution layer has been greatly improved.

Fig. 13 shows the fitting results of some of the dependent variable parameters fitted by EF and CNN-GRU. In the figure, the MAPE values of the CNN-GRU for τ, k_1 and k_2 parameters are 0.901 %, 3.47 % and 3.88 %, respectively, with significantly higher accuracy compared with EF, especially in the prediction of ignition delay coefficient (τ). Because τ has a strong regularity with the change of different working conditions, both EF and CNN-GRU neural network have a significant improvement in prediction accuracy for τ compared to k_1 and k_2, but the MAPE value of τ by CNN-GRU neural network is still 73.58 % higher than that of EF. The CNN-GRU can well fit and predict the relationship between model parameters and operating parameters, and the fitting accuracy and prediction accuracy are significantly higher than those of EF. Comparing to EF, CNN-GRU neural network does not need to consider the form of the function and fitting method, and has better ability to fit the nonlinear relationship between parameters.

5. Construction of 0-D prediction model

5.1. The construction method of engine prediction model

Based on the improved cumulative fuel mass method combined with EF or parameter relationship identification model, the pre-
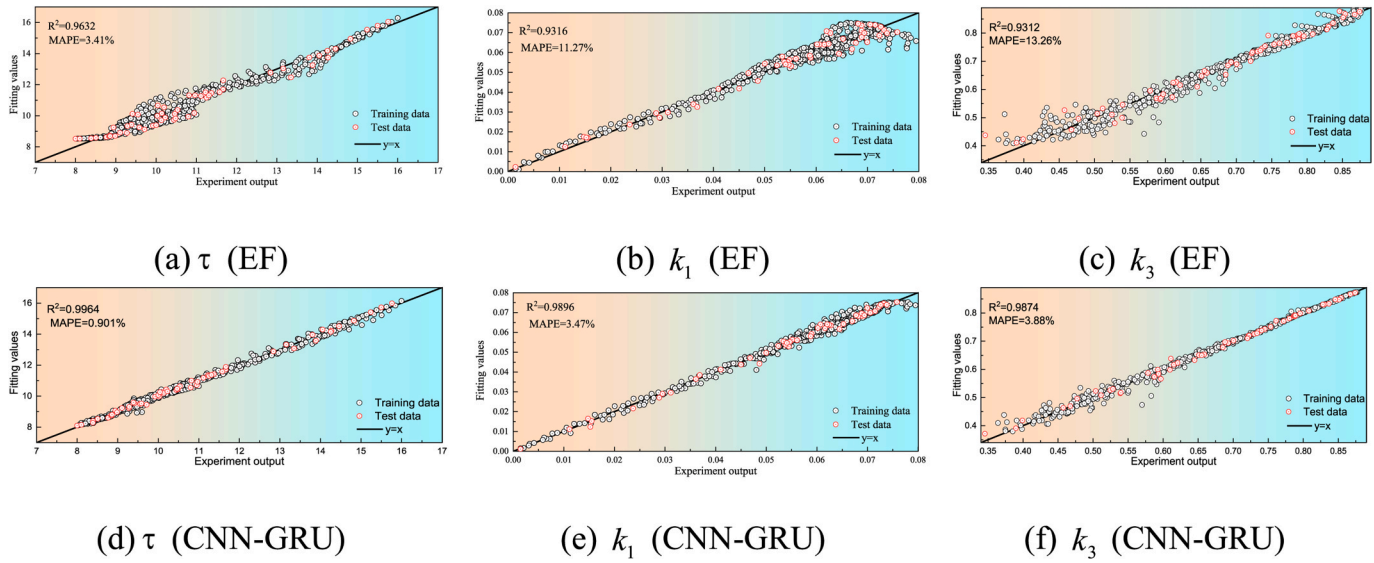
(a) τ (EF) (b) k_1 (EF) (c) k_3 (EF)

(d) τ (CNN-GRU) (e) k_1 (CNN-GRU) (f) k_3 (CNN-GRU)

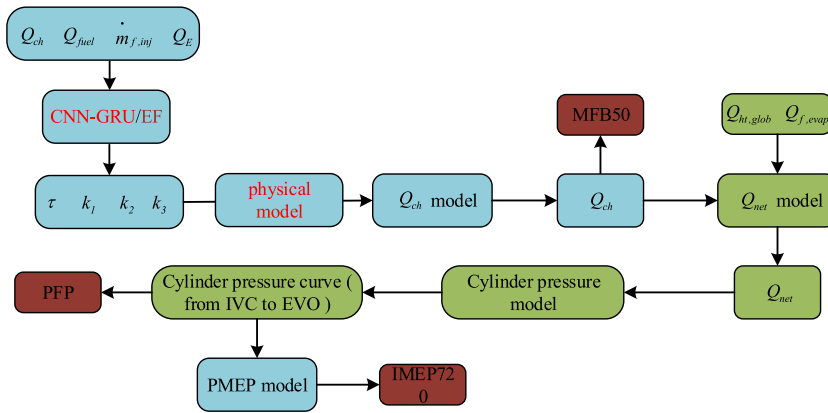**Fig. 13.** Fitting results of dependent variable parameters based on EF and CNN-GRU.

Fig. 14. The construction process of engine prediction model.

diction model can be constructed and the combustion results can be predicted. In Fig. 14, according to the operating parameters and state parameters obtained from the test, the unknown parameters in the improved cumulative fuel mass formula are calculated using EF or the parameter relationship identification model, and the construction of the combustion model (Q_{ch} model) is completed. The combustion model can calculate heat release rate, and the combustion characteristic parameter MFB50 can be extracted. According to the simplified in-cylinder heat transfer model and evaporative heat absorption model, the explicit heat release (Q_{net}) can be obtained. Then, Q_{net} is used as the input parameter of the cylinder pressure model, and the corresponding cylinder pressure curve is iterated by the Runge-Kutta method, from which the PFP can be extracted. Finally, the IMEP720 can be calculated by using the pumping loss model.

5.1.1. Reconstruction of cylinder pressure curve

From the above, the chemical heat release Q_{ch} from the combustion model based on the cumulative fuel mass approach does not consider the effects of in-cylinder heat transfer and evaporative heat absorption, so it is necessary to simplify in-cylinder heat transfer model and evaporative heat absorption model. The total heat transfer loss in the combustion process is simplified as the difference value between the maximum value of the fuel chemical energy release from ignition moment and the maximum value of the explicit heat release curve, denoted as $Q_{ht,glob}$. The expression of the explicit heat release curve is shown as Eq. (10):

$$Q_{net}^{SOC} = Q_{ch} \frac{\dot{m}_{f,inj}(t) H_L - Q_{ht,glob}}{\dot{m}_{f,inj}(t) H_L} \tag{10}$$

The fuel evaporation process runs through the two stages of compression process and combustion process. The heat absorption quantity by fuel evaporation $Q_{f,evap}$ is mainly determined by the in-cylinder temperature, the temperature of injected fuel and injected fuel quantity. The fuel evaporation process is accompanied by complex chemical reactions such as fuel decomposition and combustion, so it is difficult to calculate accurately according to the cylinder pressure curve and the burnt fraction curve. In order to facilitate the rapid estimation of $Q_{f,evap}$, the calculation of $Q_{f,evap}$ is simplified, that is, the reduction in heat quantity from the start of injection to the minimum value of the explicit heat release curve. The optimized explicit heat release curve is shown in Eq. (11).

$$Q_{net}^{SOI} = Q_{net}^{SOC} - Q_{f,evap} \tag{11}$$

$Q_{ht,glob}$ and $Q_{f,evap}$ are the same as the model parameters, which can be calibrated by the test values of specific working conditions. The in-cylinder thermodynamic process consists of five stages: intake, compression, combustion, expansion and exhaust. The combustion model based on the improved accumulated fuel mass approach can calculate the heat release rate curve during the combustion process. Then, the heat release rate curve is combined with the initial state parameters of the engine, and the cylinder pressure curve of each process is calculated by the Runge-Kutta method.

5.1.2. Simplification and reconstruction of MFB50, PFP and IMEP

The value of MFB50 can be obtained by interpolation calculation of the burnt fraction curve, as shown in Eq. (12):

$$MFB50 = \text{interp}\left(x_b, \frac{max(x_b)}{2}\right) \tag{12}$$

where x_b is the burnt fraction curve.

From Fig. 14, using Q_{net}^{SOI} as the input parameter of the cylinder pressure model, the cylinder pressure curve can be reconstructed, and the PFP can be extracted from the curve, as shown in Eq. (13):
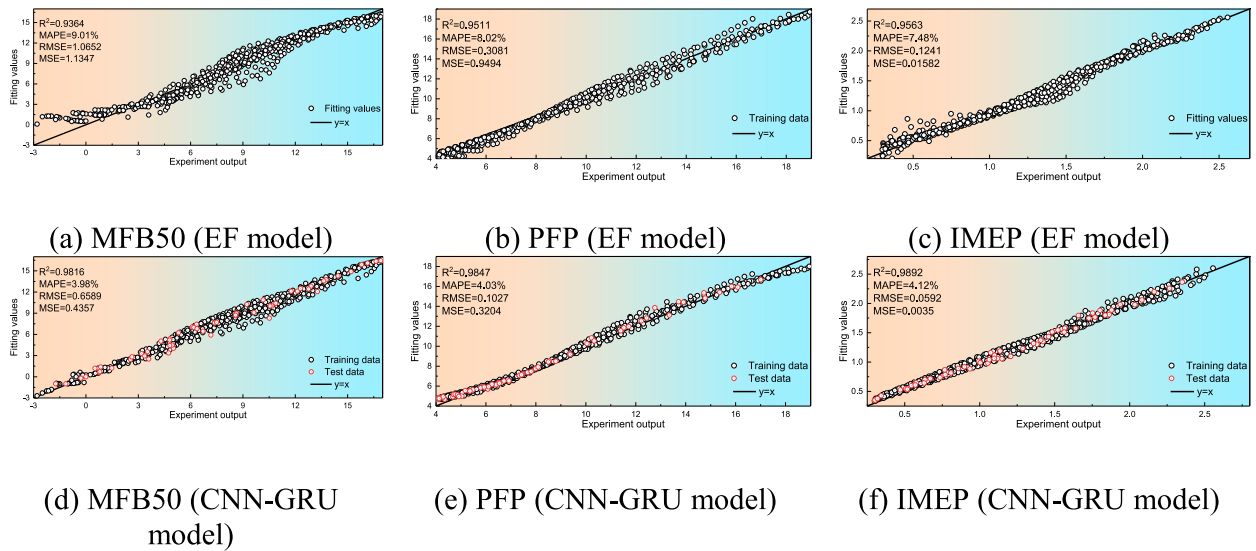
$$PFP = \text{interp}(p, max(p)) \tag{13}$$

(a) MFB50 (EF model)

(b) PFP (EF model)

(c) IMEP (EF model)

(d) MFB50 (CNN-GRU model)

(e) PFP (CNN-GRU model)

(f) IMEP (CNN-GRU model)

Fig. 15. Comparison of simulation results of 0-D prediction models based on EF and CNN-GRU under calibration conditions.

where p represents the cylinder pressure curve from inlet valve close (IVC) to exhaust valve open (EVO).

IMEP720 is the integral of the total in-cylinder pressure to the volume, as shown in Eq. (14):

$$\begin{cases} IMEP720 = \dfrac{\displaystyle\int_{180}^{540} pdV}{V_d} - PMEP \\[4mm] PMEP = \dfrac{\displaystyle\int_{-180}^{180} pdV}{V_d} \end{cases} \tag{14}$$

where V_d is cylinder clearance volume, $PMEP$ is pumping mean effective pressure, that is, pumping loss.

### 5.2. Comparison of prediction performance under calibrated conditions

A 0-D prediction model based on hybrid drive can be constructed by combining the cumulative fuel mass approach with EF or CNN-GRU neural network identification model, and the combustion results of different operating conditions can be predicted. Fig. 15 shows the comparison of simulation results of 0-D prediction models based on EF and CNN-GRU under calibration conditions. From the figure, we can observe that the prediction accuracies of the above two models are lower than those of EF and CNN-GRU neural networks identification model. The main reason is that the 0-D prediction model is based on the parameter identification model and the cumulative fuel mass approach, the accuracy will be significantly reduced due to the superposition of the two errors. The IMEP value of the 0-D prediction model based on CNN-GRU is 4.12 %, which is lower than MFB50 and PFP, but still 44.92 % higher than the IMEP value of the prediction model based on EF, and the prediction accuracy is still at a high level.

### 5.3. Comparison of prediction performance under non-calibrated conditions

To verify the prediction accuracy of the 0-D prediction models based on EF and CNN-GRU under non-calibration conditions, In Fig. 16, the comparison between the experimental and predicted values of the two prediction models under dynamic conditions. It can be seen that under non-calibration conditions, the RMSE values of IMEP (Fig. 16(b)), PFP (Fig. 16(a)) and MFB50 (Fig. 16(c)) of the model based on CNN-GRU are 0.11367, 0.5862 and 0.9824, respectively, which are 40.39 %, 29.29 % and 40.36 % lower than those of the prediction model based on EF. The above results show that under non-calibration conditions, the prediction model based on EF has low prediction accuracy and poor generalization, while based on CNN-GRU, the prediction model still has high prediction accuracy and good generalization.

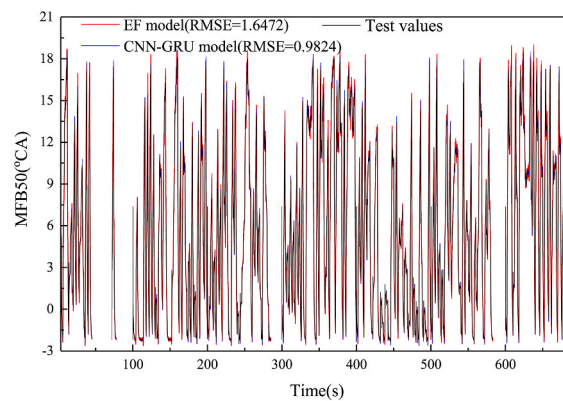### 5.4. Proposed model development flow diagram

The structure of the article is shown in Fig. 17. First, we introduced the test bench and completed the selection of test operating
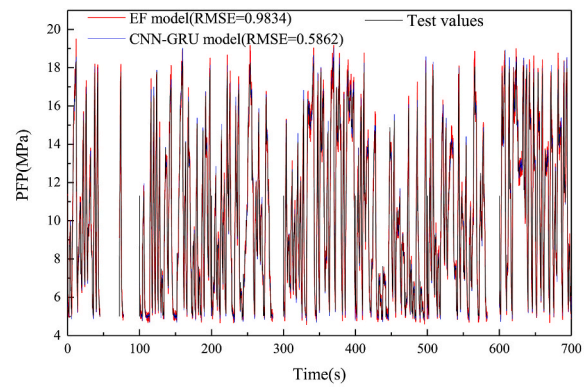
(a) PFP
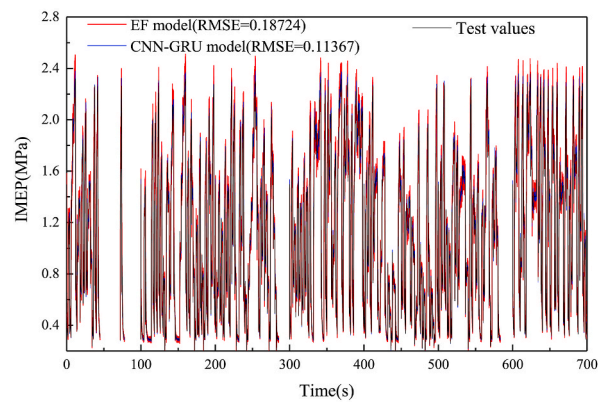
(b) IMEP

(c) MFB50

Fig. 16. The simulation results of the prediction models based on EF and CNN-GRU under uncalibrated conditions.

points. Second, based on the classical cumulative fuel mass model, the effect of in-cylinder temperature on the heat release rate of combustion process is introduced, and an improved cumulative fuel mass model is proposed. Third, the number of operating parameters is first determined by Pearson correlation coefficient combined with principal component analysis (PCA). Fourth, based on the improved combustion model, the prediction model of combustion process is constructed by combining EF and CNN-GRU respectively, and then the combustion results are simplified and reconstructed. In the final results, the advantages and disadvantages of the two prediction models based on EF and CNN-GRU are determined by comparing the prediction performance of steady-state and transient conditions.
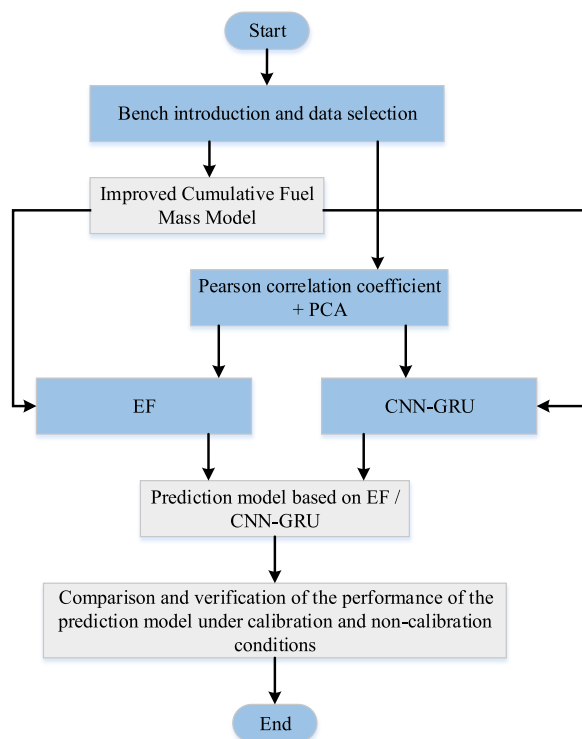
Fig. 17. Proposed CNN-GRU model development process flow diagram.

6. Conclusion

This work proposes a hybrid-driven model, represented as a prediction model based on the improved cumulative fuel mass approach combined with a data-driven CNN-GRU identification model to reconstruct the combustion process of the engine in different environments. The prediction accuracy of the model is validated.

(1) The classical prediction model based on cumulative fuel mass approach is structurally optimized. By setting different in-cylinder temperatures, the influence of in-cylinder temperature difference on heat release rate is verified. The increase of internal energy accumulated during fuel injection is introduced as an effective method to regulate the heat release rate. For the single injection engine, the optimized model has high accuracy, especially in the diffusion combustion stage where the fitting accuracy of the heat release rate is significantly improved.
(2) The advantages of Pearson correlation coefficient and PCA are combined. By analyzing the correlation between different parameters and considering the interaction between multiple factors, the dimension reduction of independent variable parameters is achieved, and the selection of independent variable set is completed.
(3) We designed a deep learning neural network including a special convolution layer and a gated neural network. CNN is used for feature identification of one-dimensional data and then embedded in GRU, which can achieve effective extraction of data features. The MAPE value of the test set of CNN-GRU neural network is 3.5 %, which is 58.38 % less than the MAPE value corresponding to EF, so the accuracy of the GRU optimized by convolutional layer is significantly improved.
(4) The prediction accuracy of the 0-D prediction model based on CNN-GRU tends to decrease significantly under non-calibration conditions, compared with the calibration condition, but it still has higher accuracy than the EF. It is more suitable as a digital twin model for real-time simulation and online optimization of engines.
(5) When calculating the heat release rate of the combustion model based on the improved cumulative fuel mass approach, the in-cylinder temperature is involved as a known parameter. However, in the actual calculation process, it is difficult to measure the in-cylinder temperature in real time and accurately. Therefore, map interpolation was performed on the measured test data for the 151 sets of temperatures, and the obtained in-cylinder temperatures were involved in the calculation. In order to further improve the prediction accuracy, the prediction model of in-cylinder temperature based on deep learning neural network can be considered.

Data availability statement

The authors do not have permission to share data.

CRediT authorship contribution statement

Deng Hu: Conceptualization, Methodology, Writing – original draft. **Hechun Wang:** Formal analysis, Software, Conceptualization. **Chuanlei Yang:** Formal analysis, Investigation, Methodology. **Binbin Wang:** Formal analysis, Investigation, Software. **Baoyin Duan:** Methodology, Writing – review & editing, Software. **Yinyan Wang:** Conceptualization, Methodology, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] H. Zhang, Y. Qin, X. Li, et al., Power management optimization in plug-in hybrid electric vehicles subject to uncertain driving cycles [J], eTransportation 3 (2020), 100029.
[2] S. Hu, H. Wang, X. Niu, et al., Automatic calibration algorithm of 0-D combustion model applied to DICI diesel engine [J], Appl. Therm. Eng. 130 (2018) 331–342.
[3] J. Castresana, G. Gabiña, L. Martin, et al., Marine diesel engine ANN modelling with multiple output for complete engine performance map [J], Fuel 319 (2022), 123873.
[4] R. Finesso, O. Marello, E. Spessa, Development of a pressure-based technique to control IMEP and MFB50 in a 3.0L diesel engine [J], Energy Proc. 148 (2018) 424–430.
[5] G. Abbaszadehmosayebi, L. Ganippa, Determination of specific heat ratio and error analysis for engine heat release calculations [J], Appl. Energy 122 (2014) 143–150.
[6] C. Ma, E.Z. Song, C. Yao, et al., Analysis of PPCI mode and multi-objective comprehensive optimization for a dual-fuel engine [J], Fuel 303 (2021), 121296.
[7] Y. Golizadeh Akhlaghi, A. Badiei, X. Zhao, et al., A constraint multi-objective evolutionary optimization of a state-of-the-art dew point cooler using digital twins [J], Energy Convers. Manag. 211 (2020), 112772.
[8] J. Guan, Y. Li, J. Liu, et al., Experimental and numerical research on the performance characteristics of OPLVCR engine based on the NSGA II algorithm using digital twins [J], Energy Convers. Manag. 236 (2021), 114052.
[9] B. Wen, X. Wu, K. Wu, et al., Ca50 estimation based on Neural Network and smooth variable structure filter [J], ISA (Instrum. Soc. Am.) Trans. 114 (2021) 499–507.
[10] Y. Li, S. Wang, X. Duan, et al., Multi-objective energy management for Atkinson cycle engine and series hybrid electric vehicle based on evolutionary NSGA-II algorithm using digital twins [J], Energy Convers. Manag. 230 (2021), 113788.
[11] Š. Kyjovský, J. Vávra, I. Bortel, et al., Drive cycle simulation of light duty mild hybrid vehicles powered by hydrogen engine [J], Int. J. Hydrogen Energy 48 (2023) 16885–16896.
[12] K.L. Ankobea-Ansah, C.M. Hall, Hybrid physics-based neural network models for diesel combustion phasing prediction [J], IFAC-PapersOnLine 55 (37) (2022) 621–626.
[13] S. Hu, H. Wang, C. Yang, et al., Burnt fraction sensitivity analysis and 0-D modelling of common rail diesel engine using Wiebe function [J], Appl. Therm. Eng. 115 (2017) 170–177.
[14] N. Sekularac, X.H. Fang, V. Shankar, et al., Development of a laminar burning velocity empirical correlation for combustion of iso-octane/ethanol blends in air [J], Fuel 307 (2022), 121880.
[15] Y. Lee, Y. Lee, S. Moon, et al., Development of semi-empirical soot emission model for a CI engine [J], Sci. Total Environ. 820 (2022), 153327.
[16] C. Mishra, P.M.V. Subbarao, Design, development and testing a hybrid control model for RCCI engine using double Wiebe function and random forest machine learning [J], Control Eng. Pract. 113 (2021), 104857.
[17] S.S. Alam, S.W. Rosa, C. Depcik, et al., Modification of the Wiebe function for methane-air and oxy-methane- based spark-ignition engines [J], Fuel 303 (2021), 121218.
[18] R. Finesso, E. Spessa, Y. Yang, Fast Estimation of Combustion Metrics in DI Diesel Engines for Control-Oriented applications[J] vol. 112, Energy Conversion and Management, 2016, pp. 254–273.
[19] F.G. Chmela, G.C. Orthabe, Rate of heat release prediction for direct injection diesel engines based on purely mixing controlled combustion [J], International Congress and Exposition2000 (2000) 152–160.
[20] A. Taheri-Garavand, A. Heidari-Maleni, T. Mesri-Gundoshmian, et al., Application of artificial neural networks for the prediction of performance and exhaust emissions in IC engine using biodiesel-diesel blends containing quantum dot based on carbon doped [J], Energy Convers. Manag. X 16 (2022), 100304.
[21] Can Ö, T. Baklacioglu, E. Özturk, et al., Artificial neural networks modeling of combustion parameters for a diesel engine fueled with biodiesel fuel [J], Energy 247 (2022), 123473.
[22] Z z Chang, M. Li, K y Zhu, et al., Model predictive control of long Transfer-line cooling process based on Back-Propagation neural network [J], Appl. Therm. Eng. 207 (2022), 118178.
[23] Y.Y. Hong, C.L.P.P. Rioflorido, A hybrid deep learning-based neural network for 24-h ahead wind power forecasting [J], Appl. Energy 250 (2019) 530–539.
[24] L. Kang, X. Zhao, J. Ma, A new neural network model for the state-of-charge estimation in the battery degradation process [J], Appl. Energy 121 (2014) 20–27.
[25] L. Hu, C. Wang, Z. Ye, et al., Estimating gaseous pollutants from bus emissions: a hybrid model based on GRU and XGBoost [J], Sci. Total Environ. 783 (2021), 146870.
[26] S. Shin, Y. Lee, Y. Lee, et al., Designing a steady-state experimental dataset for predicting transient NOx emissions of diesel engines via deep learning [J], Expert Syst. Appl. 198 (2022), 116919.
[27] S.M. Shahid, S. Ko, S. Kwon, Real-time abnormality detection and classification in diesel engine operations with convolutional neural network [J], Expert Syst. Appl. 192 (2022), 116233.
[28] G. Alcan, E. Yilmaz, M. Unel, et al., Estimating soot emission in diesel engines using gated recurrent unit networks [J], IFAC-PapersOnLine 52 (5) (2019) 544–549.
[29] Q. Shen, G. Wang, Y. Wang, et al., Prediction model for transient NOx emission of diesel engine based on, CNN-LSTM Network [J]. 16 (14) (2023) 5347.
[30] S. Hu, Finesso d'Ambrosio, et al., Comparison of Physics-Based, Semi-empirical and Neural Network-based Models for Model-based Combustion Control in a 3.0 L Diesel Engine [J] 12 (18) (2019) 3423.
[31] Z. Liu, J. Liu, Machine learning assisted analysis of an ammonia engine performance [J], J. Energy Resour. Technol. 144 (11) (2022), 112307.
[32] Y. Yan, Z. Liu, J. Liu, An evaluation of the conversion of gasoline and natural gas spark ignition engines to ammonia/hydrogen operation from the perspective of laminar flame speed [J], J. Energy Resour. Technol. 145 (1) (2022), 012302.
[33] L. Zhiyong, Z. Hongdong, Z. Ruili, et al., Fault Identification Method of Diesel Engine in Light of Pearson Correlation Coefficient Diagram and Orthogonal Vibration Signals [J], vol. 2019, 2019, pp. 1–10.

[34] M. Zhang, Y. Li, J. Cai, et al., Research on Fault Diagnosis of Diesel Engine Based on PCA-RBF Neural Network [J], 2019.
[35] S. Salam, T. Choudhary, A. Pugazhendhi, et al., A Review on Recent Progress in Computational and Empirical Studies of Compression Ignition Internal Combustion Engine, vol. 279, FUEL, 2020, 118469, 2020, 281.
[36] M.E. Flood, M.P. Connolly, M.C. Comiskey, et al., Evaluation of single and multi-feedstock biodiesel – diesel blends using GCMS and chemometric methods [J], Fuel 186 (2016) 58–67.